76-24,386

MCGRAW, Kenneth Oakley, 1944-
MONETARY REWARD AND WATER-JAR TASK PERFORMANCE: EVIDENCE OF A DETRIMENTAL EFFECT OF REWARD ON PROBLEM SOLVING.

The University of Oklahoma, Ph.D., 1976
Psychology, experimental

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

MONETARY REWARD AND WATER-JAR TASK PERFORMANCE: EVIDENCE OF A DETRIMENTAL EFFECT OF REWARD ON PROBLEM SOLVING

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial filfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

By

KENNETH O. MCGRAW

Norman, Oklahoma

1976

MONETARY REWARD AND WATER-JAR TASK PERFORMANCE:

EVIDENCE OF A DETRIMENTAL EFFECT OF REWARD ON PROBLEM SOLVING

A DISSERTATION

APPROVED FOR THE DEPARTMENT OF PSYCHOLOGY

BY

DISSERTATION COMMITTEE

ACKNOWLEDGEMENTS

I wish to thank Dr. Roger Fouts, Dr. W. Alan Nicewander, Dr. Roger Mellgren, and Dr. Lauren Wispé for the instruction they have provided over the past four years and most recently for their efforts on my behalf as dissertation committee members. Principally, however, I wish to thank Dr. John C. McCullers, my committee chairman and longtime mentor. He was the first teacher to excite me about psychology, and, ever since, he has been constantly willing to give of his own time, energy, and talent to develop in me the skills needed to do work in psychology.

For financial support over the past year I am grateful to the National Institute of Mental Health Grant No. 1-R01 MH 26359-01 PCR. Funds for the dissertation research were provided by NIMH, PHS/DHEW grant 1-R03 MH22041-01A1.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

ABSTRACT

The effect of reward with which we are all most familiar is the facilitating one so frequently found in studies that measure the rate, speed, and persistence of responding. Not so familiar are reward's detrimental effects. As the first part of the dissertation, literature on both the facilitating and detrimental effects of reward was reviewed in order to determine the crucial characteristics distinguishing tasks on which reward facilitates from tasks on which it has a detrimental effect. Two such characteristics were found. The first regards the subject's initial attitude toward the task (the aversive-attractive dimension) and the second, the type cognitive behavior required to solve the task as efficiently as possible (the algorithmic-heuristic dimension). *Post Hoc* these two dimensions account rather well for reward-nonreward differences uncovered in studies using perceptual motor, operant, verbal learning, perceptual recognition, incidental learning, discrimination learning, and problem solving tasks. Reward has been detrimental only when the task was attractive to the subject and required heuristic behavior for its solution.

The dissertation research has been to look at the effect of reward on the discovery of a "nonset" solution. Two studies were conducted. In both, college students were either monetarily rewarded or not for solving a series of water-jar problems. It was

hypothesized that reward subjects would take longer to find a simple (A-C) solution after developing a mental set to use a more complicated (B-A-2C) solution than would nonreward subjects. In both studies, poorer reward group performance was obtained as predicted. Explanations for these differences in terms of differences in either mathematical ability, time taken to check answers, or concern over quality of performance were considered and discounted. These results are consistent with the view that reward will have a detrimental effect on performance when it is superfluous to creating an initially high level of motivation and when the task is one that requires heuristic behavior for its solution.

MONETARY REWARD AND WATER-JAR TASK PERFORMANCE:
EVIDENCE OF A DETRIMENTAL EFFECT OF REWARD ON PROBLEM SOLVING

CHAPTER I

INTRODUCTION

There is a general assumption among psychologists, educators, and businessmen that the addition of external incentives to a situation will improve performance, yet there is a growing literature which can be used to argue that this is not necessarily the case. In the children's discrimination literature in particular, studies using between-group designs have shown that children who are rewarded for correct responses (M&Ms, marbles, tokens, and money are typically used) do not learn as quickly as ones who are given only knowledge of results. A host of studies supports this conclusion: McCullers and Martin, 1971; McGraw and McCullers, 1974; Miller and Estes, 1961; Schere, 1969; Spence, 1970; Spence and Dunton, 1967; Spence and Segner, 1967; and Terrell, Durkin, and Wiesley, 1959.

But laboratory discrimination studies like these in which reward is provided trial-by-trial do not on the surface at least bear much resemblance to the type performance situations educators, businessmen, and mothers are concerned with. The question becomes, then, whether the poorer performance by reward groups in discrimination learning is a trivial case, a phenomenon complete in itself with no wider implications for the effects of reward in general, or whether the poorer performance observed in this one instance is an outcropping of a heretofore unrecognized detrimental effect which reward can have on behavior.

A review of the literature indicates that the detrimental effect of reward may indeed be more general than the discrimination learning results alone would imply. In adult incidental learning, for example, reward has been frequently observed to hinder performance on the incidental task, the one requiring use of previously irrelevant stimuli. This has been the case both when the incidental stimuli were spatially removed from (Bahrick, Fitts, and Rankin, 1954; McNamara and Fisch, 1964; Johnson and Thomson, 1960; Davis and Lovelace, 1963) and spatially contiguous with (Bahrick, 1952; Staat and McCullers, Note 4) the stimuli of the intentional task.

A detrimental effect of reward on performance has been observed on other tasks outside the discrimination learning paradigm as well. Kruglanski, Friedman, and Zeevi (1971), for example, found that fifth-grade school children offered an extrinsic incentive for performance on a battery of five tests designed to measure recall and creativity were inferior in both recall and creativity to a group that was offered no incentive. Weick (1964), using a series of concept attainment tasks, found the same effect with adults. The subjects rewarded with an external incentive found fewer concepts in a 20-minute experimental session, expended less effort, and rated the task as less interesting than did nonrewarded subjects. In both the above studies, reward was nonmaterial and noncontingent indicating that the detrimental effect of reward may not be restricted just to the contingent, material rewards employed in the children's discrimination literature.

Condry (Note 1) and McGraw and McCullers (Note 3) have used concept attainment tasks and monetary incentives to obtain further evidence for a detrimental effect of reward. The subjects in the former study were high school students. The reward subjects among them received $.50 for each of six concept attainment problems they solved. On two measures of performance--number of information requests needed to reach solution and proportion of logical to illogical requests--reward subjects were shown to be inferior to nonreward subjects. McGraw and McCullers obtained similar results using adults. In this study the time taken to achieve a solution was recorded along with the number of information requests. On each of four problems, reward subjects ($.50 per solution) took longer and asked more questions than did nonreward subjects, but the differences were not statistically reliable.

Adult problem solving studies offer few instances of reward -nonreward comparisons, but those that exist tend to show reward to have a detrimental effect. Viesti (1971), for example, describes two experiments he conducted using an "insight learning task." On each page of a fifteen page booklet there appeared a set of three computer generated patterns. Two of the patterns were 2/3 redundant and the other was completely different. The subject's job was to say which of the three patterns was different. In both studies, nonreward subjects were correct more often than reward subjects. Two studies by Glucksberg (1962, 1964) have shown reward subjects take longer to solve difficult functional-fixedness problems than nonreward subjects.

In summary, there is substantial evidence to encourage the view that reward may have negative consequences on performance that are of general importance. The evidence is extensive in terms of subject populations sampled, tasks, procedures, and dependent measures. More than the results of any one study, it is this breadth of instances in which reward has been observed to produce a detrimental effect which makes it appear unlikely that the reward-nonreward differences in children's discrimination learning are artifactual.

Alongside the literature that shows reward to have a detrimental effect, there is the more voluminous and familiar literature which shows reward to be a facilitator. Elsewhere (Note 2) I have contrasted studies which show reward to have a detrimental effect with those that show reward to facilitate on the basis of two dimensions which seem to be crucial to determining what reward's effect will be in a specific case. These were named the aversive-attractive and algorithmic-heuristic dimensions.

The aversive-attractive dimension specifies the subject's attitude toward the task. Aversive tasks are those on which the subject would not choose to participate were it not for some incentive. Examples from standard laboratory tasks are those which require sustained and unchallenging perceptual-motor responding as in the various lever-pulling, vigilance, and tracking tasks. In clear contrast with these are tasks which are inherently interesting such as concept formation and other types of problem-solving tasks. For these, the offer of an incentive is superfluous since incentives are not needed to obtain an initially high level of motivation. One characteristic shared by all the tasks used in studies showing reward to have a detrimental effect is that they have been attractive.

While attractiveness may be a necessary attribute of tasks which show reward to have a detrimental effect, it is not sufficient because there are numerous examples of reward facilitation on attractive tasks. A good example of this is speed tests of mental multiplication (Weinstein, 1971a, 1971b, 1972). For this reason it is necessary to specify the algorithmic-heuristic dimension as second means of distinguishing tasks on which reward will facilitate from those on which it will have a detrimental effect.

Mental multiplication was chosen as an example of an attractive task on which reward facilitates performance because what is meant by algorithmic is clearest in this case. "Algorithmic" problems such as mental multiplication may be difficult, but there is no question about where to begin or how to proceed. In contrast with mental multiplication problems are heuristic problems, such as those requiring insight and creativity, in which the whole problem is determining what steps are relevant for a solution. This characteristic, like attractiveness, has been common to all the studies in which reward has had a detrimental effect.

A task which lies near the attractive and heuristic extremes of the aversive-attractive and algorithmic-heuristic dimensions is the mental set task employing water-jar problems that was used by Luchins (1942). This task is designed to establish a mental set for a relatively complicated, indirect solution to a type of arithmetic problem and then to test whether a subject is subsequently able to find a simpler, direct solution to the same type problem. Such a task is heuristic because breaking a mental set is paradigmatic of what is meant by heuristic behavior. The attractiveness of the task for college students, who tend to view it as a test of intelligence, is evident to any one who has ever administered it. On the basis of the model developed from the available literature then, a water-jar task should be optimal for producing a detrimental effect of reward on performance. Consequently, it was chosen for use in the present research, the goal of which was to add to the literature on the detrimental effect of reward. A larger literature demonstrating this effect was deemed necessary for two reasons. First, it is not commonly acknowledged at present that there is any such thing as a "detrimental" effect of reward on performance. Second, if theorizing in this area is to be fruitful, a larger literature than is currently available needs to accumulate so that speculation on the causes of reward's detrimental effect will not be based on the results of studies from a single area, studies that are methodology bound, or unreplicated studies.

CHAPTER II

EXPERIMENT I

In the first experiment, adult subjects were given a series of ten water-jar problems. Problems one to five were soluble in a single way by a well-defined rule which involved using all three of the available jars in a precise way. The next four problems could be solved using the same solution as problems one to five (the mental set solution) but there was, in addition, a direct solution that required using only two of the jars. This solution was conceptually and mathematically simpler and an obvious solution if subjects could but see it. Problem ten in the series--which after Luchins (1942) we call the extinction problem--was designed to force subjects to see the alternative, two-jar solution since no three-jar combination would provide the correct answer. Given Luchins' (1942) discovery of the persistence with which mental set solutions are applied, it was expected that most subjects would not find the novel solution until the last problem, so time-to-solution on this problem provided a convenient measure of the facility with which subjects could break set. The specific predictions made for performance were that subjects offered a reward for each correct solution would, first, be less likely to find the direct solution on problems six to nine and, second, take longer to solve the extinction problem than would nonreward subjects.

Method

Subjects and Design.

The subjects were thirty undergraduates enrolled in introductory psychology at the University of Oklahoma whose participation in the experiment partially fulfilled a course requirement. Subjects were randomly assigned by pairs to the reward and nonreward experimental conditions with the first pair member going to one group and the second to the other. There were thus fifteen subjects per group. The design of the study was a simple groups-by-trials design with two levels of the between factor (reward and nonreward) and ten levels of the within factor (problems).

Task.

The ten water-jar problems that made up the task are presented in Table 1. The first five problems could only be solved using the formula B-A-2C meaning that jar "B" was to be filled, and then, from "B", jar "A" was to be filled once and jar "C" twice, leaving the desired amount as a remainder in "B". Problems six through nine had dual solutions. The subjects could either continue to use the old formula (B-A-2C) or switch to an alternative procedure (either A-C or A+C depending on the problem). The last problem in the series was the only one not soluble by the formula B-A-2C. Its solution was A-C.

Procedure.

Five sample problems were used to introduce the subject to the task and to familiarize him with the method of recording his

Table 1

Water Jar task for Experiment 1

Number & Type of Problem		Given these jars A	B	C	Obtain this amount
1.		19	139	9	102
2.	Single-solution	14	163	25	99
3.	Single-solution	18	43	10	5
4.	Single-solution	16	79	11	41
5.		9	64	15	25
6.	Dual-solution	23	49	3	20
7.	Dual-solution	15	39	3	18
8.	Dual-solution	14	36	8	6
9.	Dual-solution	18	48	4	22
10.	Extinction problem	27	80	4	23

Water Jar task for Experiment 2

The following problems were substituted for 6,7,8, and 9 above.

		A	B	C	Obtain this amount
6.	Dual-solution	18	125	5	97
7.	Dual-solution	13	42	9	11
8.	Dual-solution	27	95	12	44
9.	Dual-solution	14	37	8	7

answer which was to write the solution as a formula, e.g., B-A-2C, A+C, A-C. The problems were typed on index cards so that the subject would work on but one at a time. Separate sheets were provided for the answers. On completing each answer, the subject put both the answer sheet and the problem card in a box to his right before going on to the next problem in the series. This procedure not only assured that the subject would not return to a problem already completed but also it provided a cue to the experimenter, who was behind a one-way glass keeping times-to-solution, that the subject had finished the problem and was ready for the next.

In the instructions the subject was told that he would have as long as he needed to find a correct solution to each of the problems and that he would be left alone to work. Reward subjects were told that they would receive a nickel for each correct answer and a fifty-cent bonus if all were correct. The bonus was provided to insure the importance of the reward on problem ten. Had a constant reward of ten cents been offered for each problem, the additional monetary inducement for solving problem ten would have been small relative to the ninety cents subjects would have already won. The rationale given for offering money at all was that the experimenter had found it necessary to "motivate people a little to get them to work until they find the correct answer" and that the little money spent was well worth the time and energy saved by not having to run extra subjects.

On completing the instructions, the experimenter went to an adjoining room from where he could observe by one-way mirror. The mirror was to the side and behind the subject who sat at a desk facing a blank wall. There was no attempt to deceive subjects regarding the mirror. Each was told that it was there and that the experimenter would look through from time to time to see how the subject was coming. The only deception used was in specifically instructing subjects that time was not crucial when in fact it was the all important dependent measure. The reason for not telling subjects that times were being recorded was to make the situation as stress free as possible. Despite their being instructed that time was not important, no subject stopped in his task to gaze about the room or to fiddle with non-task materials. Without exception, subjects moved from one problem to the next in a very conscientious manner.

Time-to-solution was recorded for each problem in the ten-problem series by means of stopwatches. Timing began when the subject turned over a problem card and ended when he placed that card along with his answer in the box to his right. When a subject had completed all ten problems, the experimenter explained the purpose of the study, answered any questions, and paid the reward subjects.

Results and Discussion

Subjects differed in when and whether they broke set. Some found the direct two-jar solution between problems six and nine; others found it only on the extinction problem (problem ten) when it was the single available solution. In addition, there were some subjects who persisted even on the extinction problem to use, erroneously, a three-jar, set solution. On the basis of such differences, subjects were divided into four categories.

Category I included subjects who used a direct two-jar solution at least twice before problem ten and then did not revert to the old three-jar solution. These subjects, on the basis of their rapid solution times for the latter problems in the series and on the basis of their post-experimental verbal reports, had clearly broken set prior to problem ten, and so did not begin work on the final problem with the expectation that the old solution would work. Category II subjects used a direct, two-jar solution on one of the problems six to eight but only on one of these and not on nine. The fact that these subjects reverted to the old solution might indicate that their one experience with two-jar solutions had not been sufficient to break their mental set.[2] Category III subjects used the B-A-2C solution in every case where it was applicable but then succeeded in finding the direct, two-jar solution on the extinction problem. Category IV subjects never broke set and after much effort gave some very complex but incorrect three-jar solution as their answer to problem ten. These four subject categories are ordered by the extent to which subjects in each are presumed to have been affected by their mental sets with Category I subjects being the least, and Category IV subjects, the most affected. A summary of the four categories and the number of subjects in each is given in Table 2. Three subjects (two nonreward and one reward) are not included in Table 2 because they made a number of errors on the first five problems and, therefore, failed to establish a mental set. Since the

Table 2

Numbers of Subjects and MeanTime -To-Solution on Problem Ten for Each of

Four Subject Categories

Categories	Reward		Nonreward	
	n	Mean time (sec.)	n	Mean time (sec.)
I. Broke set on two or more problems prior to Ten	2	25.8	3	25.3
II. Broke set on only one problem prior to Ten and then reverted to B - A - 2C	2	205.1	3	44.9
III. Broke set only on Problem Ten	7	236.3	6	134.7
IV. Never broke set	3	431.6[a]	1	431.6[a]
Totals	14[b]		13[b]	

[a]These subjects failed to solve Problem Ten and so were assigned a solution time equal to the slowest time of the Category III subjects.

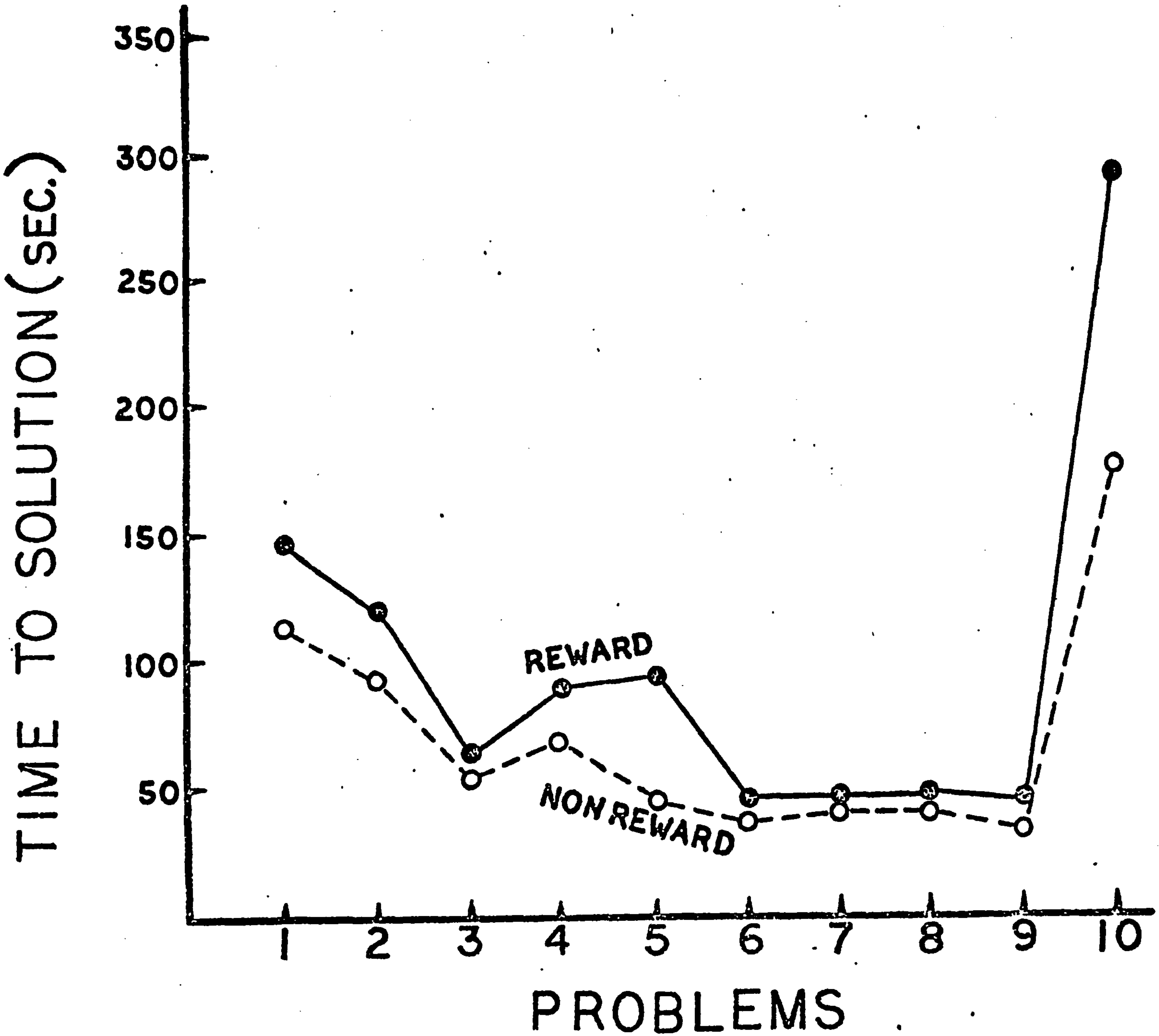
[b]One reward and two nonreward subjects were eliminated for failing to establish set.

hypotheses to be investigated regarded differential performance in breaking set, the data of these subjects who never established set became irrelevant and were eliminated from the principal analyses.

Due jointly to the proliferation of categories, the loss of the three subjects who failed to establish set, and the initially small sample, a statistical test of the hypothesis that nonreward subjects would break set earlier in the ten problem series than the reward subjects was impracticable for lack of power. Nevertheless, since the subjects in Categories I and II are those who were least blinded and those in Category IV the subjects who were most blinded by their mental sets, it is in conformity with expectations that there were more nonreward subjects in the former two categories and fewer in the latter than reward subjects (6 to 4 and 1 to 3 respectively).

In addition to the prediction that nonreward subjects would be more likely to find a direct solution prior to problem ten, it was predicted that nonreward subjects would be faster to find the A-C solution on the extinction problem than would the reward subjects. The mean extinction problem solution times for the reward and nonreward groups are given by subject category in Table 2. Category IV subjects, because they failed to solve problem ten, had to be assigned a solution time for this problem. On the basis of the judgment that Category IV subjects were inferior in performance to subjects in Category III, the four subjects in Category IV were assigned the same solution time as was obtained by the slowest of the Category III subjects, 431 sec. Using this procedure, the mean solution time for the nonreward subjects on problem ten was 112 sec. and for the reward subjects it was 244 sec., $t(25) = 2.36$, $p < .05$. This difference clearly supports the prediction that nonreward subjects would be faster to solve the extinction problem. Figure 1 shows that this superiority of nonreward subjects, which was dramatically present on problem ten, was to some extent present on the other nine problems as well. A group by trials analysis of variance revealed this consistent difference to be significant, $F(1,15) = 5.65$, $p < .05$.

The possibility was considered that the solution-time differences on problem ten might simply indicate that reward subjects took greater care with their answers than did nonreward subjects. One argument against this possibility is that the problems required only simple addition and subtraction. Because of this, it is logical to assume that most of the time recorded here as "solution time" was in fact time spent to discover and not to check an answer. This argument is particularly applicable to problem ten where it is difficult to imagine subjects spending appreciable time to verify the result 27 - 4 = 23. Observations support this view. On problem ten, the modal subject gave some visible sign of discovery. The time taken to record and turn in an answer following this was never more than about 15 seconds. Further supporting the view that it was discovery time and not checking time which differentiated reward from nonreward subjects on problem ten are the data for subjects in Category I (Table 2). For these subjects who had already discovered the direct solution prior to problem ten, the time taken to verify, record, and turn in their answers was a uniform 25 seconds.



TIME TO SOLUTION (SEC.)
350
300
250
200
150
100
50
REWARD
NON REWARD
1
2
3
4
5
6
7
8
9
10
PROBLEMS

A complement to the argument that reward subjects took more time because they took more care is the argument that nonreward subjects took less time because, having no need to be overly concerned whether an answer was correct or incorrect, these subjects emphasized speed at the expense of quality. Had this been the case, there should have been some indications of poorer quality of performance by nonreward subjects in a measure such as number of errors.

In designing the task, care was taken to make errors minimally probable by giving subjects a number of practice problems, unlimited time to work the task problems, and ample scratch paper to verify each result. Despite these precautions, errors were still made. Solutions judged to be errors included not only miscalculations but also unrecorded answers (caused by a subject's giving up), and solutions which were mathematically correct but in violation of the terms of the problem. Illustrating such an error is the solution given by one subject that depended on the presence of a negative three units of water in Jar B following one of the operations. There were five errors of this latter sort. Responsible for four of them were the two nonreward subjects who were eliminated for failing to establish set. Out of the 150 solutions (10 problems × 15 subjects) from each group, there were a total of fifteen errors made by reward subjects and nine by nonreward subjects. Contributing to these totals were eight of the reward subjects as compared to just four of the nonreward subjects. From these results it can hardly be argued that nonreward subjects acheived their speed of responding at the expense of quality. To the contrary, it is the quality of the reward group performance which appears to be inferior. This is a surprising result in light of the fact that reward subjects were specifically told that obtaining the maximum reward of $1.00 was contingent on their finding a correct solution for each of the ten problems.

It has been argued that the faster nonreward solution times on problem ten can be attributed neither to differences in quality of performance nor to differences in time used to check answers. The question remains whether the nonreward superiority can be attributed solely to the reward variable. Since it would seem that mathematically ability should play a large part in how well a subject performed on the water-jar task, we sought *post hoc* to determine whether the groups were evenly matched on this characteristic. As estimates of mathematical ability, college-bound percentiles based on scores on the mathematics usage section of the American College Test (ACT) were available for eleven reward and eleven nonreward subjects. This twenty-two-subject subset of the twenty-seven subjects included in the analysis differed significantly in aptitude, $\underline{t}(20) = 2.14$, $\underline{p} < .05$. The reward subjects had the lower scores meaning that mathematical aptitude was partially confounded with treatment. Dividing ACT scores at the median revealed that seven of the eleven subjects below the median were in the reward group. With the mean ACT percentile being .67, the mean for reward subjects was .57 and for nonreward subjects, .78. The sample estimate of the correlation of ACT percentiles with solution times on problem ten, $\underline{r} = -.32$, $\underline{p} > .10$, gives an indication of the magnitude of the effect this confounding might have had on the obtained group differences. While nonsignificant, this value is large enough that because of the partial confounding with treatment, it poses some problem for the interpretation of the data offered here. Since there is no way to correct for confounding after the fact, the question of whether the reward variable would have a significant effect even when not partially confounded with mathematical ability was left for the second experiment.

CHAPTER III

EXPERIMENT II

Because the reward group in the previous study had a disproportionate number of low ability subjects, we undertook a replication prior to reporting our findings.

In the replication, a few procedural changes were made. The dual solution problems six through nine from the water-jar task were replaced by single-solution set problems. Also, subjects who recorded incorrect answers to problem ten were asked to continue until they found a correct solution. These two changes were designed to eliminate subject categories I, II, and IV, that is, subjects who either broke set early or not at all. This simplified the test of the hypothesis that reward subjects would take longer to discover the A-C solution on the extinction problem than would nonreward subjects. A third change was to increase from fifty cents to one dollar the bonus for getting all ten problems correct. This raised the total reward from \$1.00 to \$1.50. The one design change was to include ACT mathematics usage percentile scores as a blocking variable. The advisability of this change was made evident by the fairly large correlation found in the first study between ACT and solution time on problem ten.

In this second study, data were also collected on subjects' attitudes toward the task. This provided both a means for verifying that the water-jar task was truly perceived as attractive by the college student subjects--we only assumed this in Experiment I--and a means for replicating work by Deci (1971), Lepper, Greene, and Nisbett (1973), Kruglanski, Alon, and Lewis (1972), and Kruglanski, Friedman, and Zeevi (1971) concerning reward's effect on motivation. This work argues that when a subject is rewarded for a task which he is already intrinsically motivated to perform, he will be induced to attribute his participation to the reward with the result that his intrinsic motivation is reduced. In the terms developed here, this would mean that reward turns attractive tasks into aversive ones. Deci (1971) and Lepper *et al.* (1973) were able to show such an effect using behavioral measures; Kruglanski and his colleagues have done the same using a questionnaire. In related work, Benson (1970) used semantic differentials to assess changes in attitude toward reading after a semester-long reward program had been used to improve reading skills. While the reward group in Benson's study held a slightly more favorable attitude than controls at the start of the study, their attitude toward reading at the end was significantly below that of controls. For convenience, semantic differentials were used in the present study. These were administered at the end of the task and again about two weeks later in order that attitude change could be studied as a function of time as well as treatment.

Method

Subjects.

The subjects were forty introductory psychology students as described in Experiment I. ACT scores were available for thirty-two of these subjects. Data were nontheless collected from the eight subjects not having ACT scores. Six additional subjects began the study but were dropped. Five were dropped because of their exceptionally poor performance on the sample problems; the sixth was dropped when he failed to find the solution to problem one within the 45 minutes allowed for the experimental session. Of these six subjects eliminated for poor performance, three would have been in the reward group and three in the nonreward group. Of the thirty-two primary subjects, sixteen were in each treatment group. A median split was used to divide subjects into high and low categories on the ACT variable. Using this procedure there were eight highs and eight lows in the reward group, seven highs and nine lows in the nonreward group.

A major purpose of this replication study was to control for mathematical ability as estimated by subjects' ACT percentiles. The sample distribution of ACT percentiles had a mean of .64 and a standard deviation of .27. The reward mean was .60 and the nonreward mean, .67. These means do not differ as indicated by the very small t value, $t(30) = .70$, $p = .50$. ACT scores, therefore, were free of any confounding with incentive condition.

Procedure.

Except for the changes noted in the introduction, Experiment II was conducted in the same way as Experiment I. Following the experimental session and after subjects had been debriefed, the first of the semantic differentials was administered. On it there were two open-ended statements: "The water-jar task was..." and "The research being done using the water-jar task is ...." Each was followed by a set of seven-point scales which had evaluative adjectives as their anchors. The choices in the set of scales following the first statement were exciting-dull, bad-good, meaninful-meaningless, inferior-superior, interesting-boring, unimportant-important, and pleasant-unpleasant. The choices following the second statement were the same except that valuable-worthless was used in place of pleasant-unpleasant. Within two to three weeks after subjects had participated in Experiment II, they were mailed a second semantic differential identical to the first along with a letter explaining that it had become necessary to collect additional information.

Results and Discussion

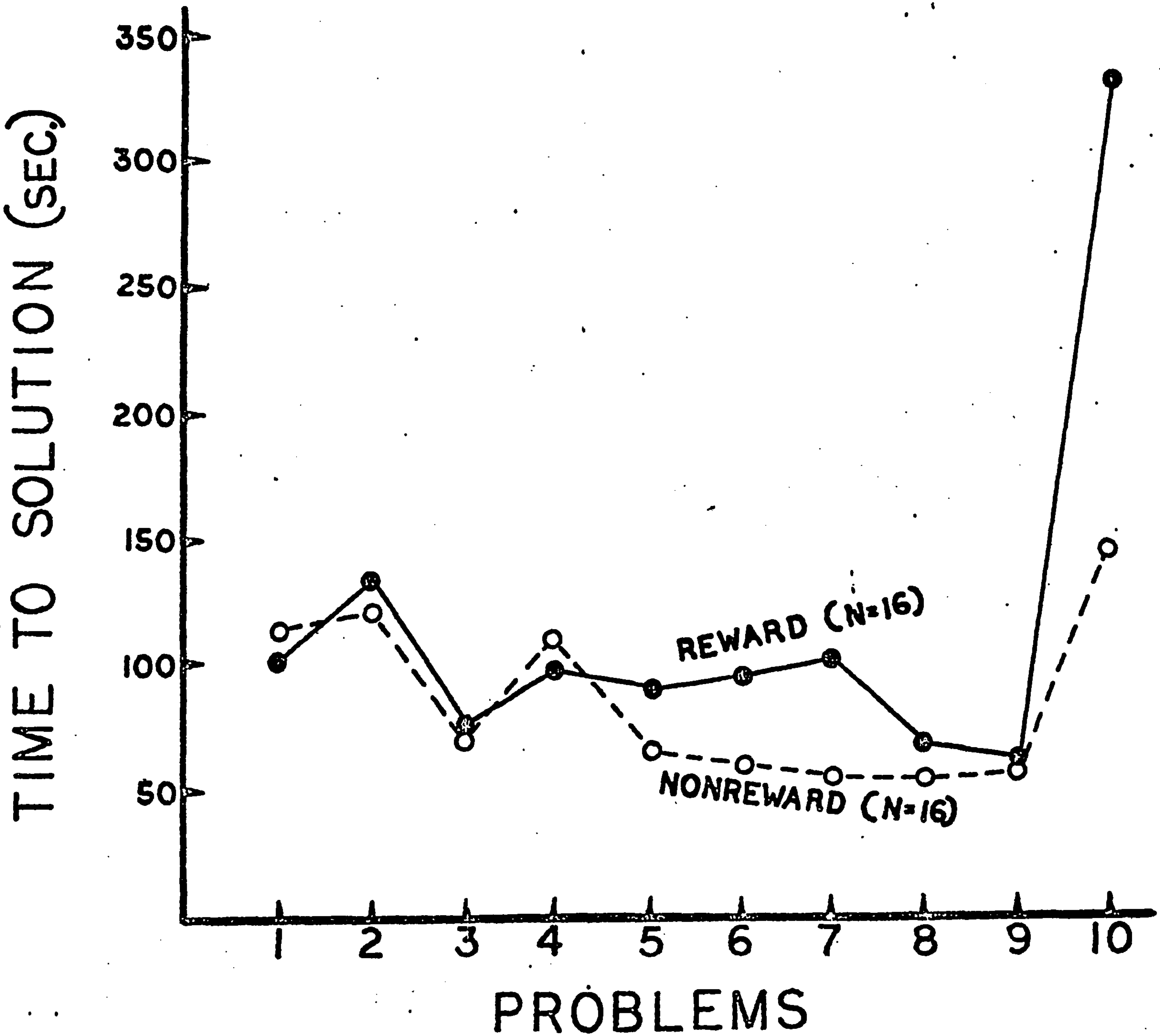
Although subjects who made initial errors on problem ten were encouraged to continue until they found the correct answer, three subjects gave up without ever discovering the A-C solution. All three were in the reward group. In order that their data could be included in the analyses, these subjects were assigned a solution time of 15 minutes. The actual times they spent attempting to find a solution were 18, 28, and 40 minutes respectively.

Solution time.

Considering just the data of those thirty-two subjects who had ACT scores, the mean solution times for the extinction problem were 143 sec. in the nonreward group and 338 sec. in the reward group, $F(1,28) = 3.48$, $p.<.07$. These results of course replicate, under conditions in which mathematical ability was controlled, the result of principal interest in Experiment I. Including the data of those subjects who had no ACT score only enhances the difference: nonreward mean = 122 sec., reward mean = 326 sec., $F\ (1,38) = 5.63$, $p<.025$.

Figure 2 graphically presents reward and nonreward solution times for each problem for the subjects who had ACT scores. As was true in the first study, there was a consistent tendency for nonreward subjects to solve the water-jar problems more quickly. This consistency coupled with the large differences on problem ten made for a significant Incentive main effect in the repeated measures analysis of variance, $F(1,28) = 6.30$, $p<.05$. Including all forty subjects in the analysis did not alter the effect, $F(1,38) = 4.71$, $p<.05$.

As in the first study, neither observations nor logic support the view that any appreciable amount of the time that separated reward from nonreward groups on the final problem (which was over three minutes in this study and over two minutes in the first) could have been taken up by activities designed to insure the correctness of the $27-4 = 23$ solution. Furthermore, it was reported in the first


TIME TO SOLUTION (SEC.)
350
300
250
200
150
100
50
REWARD (N=16)
NONREWARD (N=16)
1
2
3
4
5
6
7
8
9
10
PROBLEMS

study that reward subjects made more errors than nonreward subjects, a fact which supports the belief that reward performance overall was inferior to that of the nonreward subjects. The same thing was true in the present study. There were eight errors made by reward subjects compared to two by nonreward subjects. Also impressive is the result that seven of twenty reward subjects as compared to only two of the twenty nonreward subjects contributed to these totals. Were reward subjects taking greater care to check their answers, it would be difficult to explain why they made so many more errors.

The results again showed the importance of mathematical ability as a predictor of solution time. The correlation of ACT percentiles with solution times on problem ten was even larger than in Experiment I, $\underline{r} = =.41$, $\underline{p}<.05$. The importance of math ability as a determinant of solution time independent of the problem number is indicated by the highly significant ACT main effect, in the repeated measures analysis, $\underline{F}(1,28) = 16.27$, $\underline{p}<.001$.

Attitude data.

The initial post-experimental semantic differential which inquired into subjects' attitudes toward the water-jar task produced the expected result that subjects generally view the task as attractive. Scores on the task-attitude semantic differential could range from a low of 7 to a high of 49. The observed range was from 21 to 49 with the mean score being 37.67. This means that on the seven 7-point evaluative scales, the average rating was 5.38. More interesting, however, than knowing that college students subjects tend to find the water jar problems attractive is knowing whether subject performance in the reward group can be patially predicted by knowing the subject's attitude toward the task. That is, if attractiveness is indeed one of the dimensions which determines when reward will facilitate and when it will disrupt performance, there should be a positive correlation in the reward group between the rated attractiveness of the task and the time it took to find the solution to problem ten. Data from reward subjects in the present study, however, show absolutely no relation, $r = .02$. Lack of variability in the attractiveness measure could well account for this nonsignificant result. For an adequate test, one would need attitude scores which range from aversive to attractive rather than have scores which vary only at the attractive end of the scale as in the present study.

To determine whether attitudes toward the task in particular and the research in general were negatively affected by reward and whether there were differences as a function of time, data from the semantic differentials administered both immediately after and two weeks after the experiment were used in groups-by-trials analyses of variance. There were two of these: one for the subject's attitude toward the task and one for his attitude toward the research in which he had participated. In neither of these analyses were there any significant effects. Attitude as measured by semantic differentials did not differ by group and did not change over time nor was there any interaction. For the task-attitude scales, the reward mean went from 38.17 on the first to 37.55 on the second; the nonreward mean went from 38.53 on the first to 37.71 on the second. For the research-attitude scale, the reward mean went from 41.67 on the first to 39.22 on the second; the nonreward mean went from 40.88 on the first to 40.41 on the second.

CHAPTER IV

GENERAL DISCUSSION

Because the principal purpose of the present study was to add to the literature on the detrimental effects of reward, we selected water jar problems for our task because they had not been previously used to investigate the effects of incentives and because problem ten of the series seemed to share two crucial characteristics with other tasks that have been used to produce a detrimental effect. These were attractiveness and a "blocked" solution requiring heuristic behavior. Discovery that reward did disrupt performance on problem ten of the water-jar series lends some credence to the importance of these characteristics but, of course, the present study in no way provided a test of their importance as boundary conditions.

The evidence in the present study for a detrimental effect of reward along with that from other studies cited in the introduction raises the question of why the assumption that reward is facilitating should be so generally held. A possible answer comes from the fact that we have largely limited our research into reward's effects on performance to tasks that are unattractive and have algorithmic solutions. Reward _is_ facilitating if one looks only at the latency, rate, or persistence of some well-honed, and therefore generally unattractive, act. But even given the limitations on our selection of tasks and dependent measures for use in research on reward's effects it is puzzling that reward should be so widely viewed as capable of improving but not hindering performance since theoretical predictions for detrimental effects or reward have been with us for some time. Yerkes and Dodson published their well-known work in 1909 and Spence did so in 1956.

One might think that given the work of Spence (1956) and Yerkes-Dodson (1908) nothing more needed to be said on the subject. This, however, is not true. First, Spence's extension of Hullian theory to include incentive effects is not compatible with the Yerkes-Dodson law in that, according to the latter, reward could have facilitating and detrimental effect on performance at a single task depending upon fluctuations in the level of motivation. This would not be predicted, however, from the Hull-Spence equation, $\underline{E} = \underline{H}(\underline{D}+\underline{K})$. Second, while the Yerkes-Dodson law has been frequently invoked by investigators manipulating a variety of motivational factors, planned studies (e.g. Broadhurst, 1959) have generally used aversive stimulation as the means of raising motivational level to the point where it disrupts performance. To produce a detrimental effect using reward it has been sufficient to offer small amounts of money to adults or candy to children and on some occasions this offer has even been noncontingent. In the present study, for example, reward subjects were under no time pressure and thus were assured of the maximum reward if they but did their arithmetic properly. As a final illustration of why the case for the detrimental effects of reward is not closed, we need only to note that it is not at all clear how reward's effect on the production of responses initially low in the habit hierarchy serves as an explanation for the occurrence of detrimental effects in discrimination learning and concept formation where competing responses should have equal rather than differential strengths at the start of training. Incidental learning poses a special problem since the detrimental effect occurs on learning for which no incentive is offered.

REFERENCE NOTES

1. Condry, J. The role of initial interest and task performance on intrinsic motivation. In J. C. McCullers, (Chair), *The Hidden Costs of Rewards*. Symposium presented at the meeting of the American Psychological Association, Chicago, 1975. (Available from the author, Department of Human Development, Cornell University, Ithaca, New York, 14850).

2. McGraw, K. O. The detrimental effects of reward on performance. In D. Greene and M. R. Lepper, (Eds.), *The Hidden Costs of Reward*, Hillsdale, N. J.: Lawrence Erlbaum Associates, book in preparation.

3. McGraw, K. O. and McCullers, J. C. Monetary reward and concept attainment: Further research on the detrimental effect of reward. Unpublished manuscript, 1975. (Available from either author, Department of Psychology, University of Oklahoma, Norman, Oklahoma 73069).

4. Staat, J. and McCullers, J. C. The distracting effect of material reward on the recall of incidentally acquired R-S associations. Paper presented at the meeting of the American Psychological Association, New Orleans, August, 1974. (Available from either author, Department of Psychology, University of Oklahoma, Norman, Oklahoma, 73069.)

LIST OF REFERENCES

Bahrick, H. P. Incidental learning under two incentive conditions. *Journal of Experimental Psychology 47,* 170-172 (1954).

Bahrick, H. P., Fitts, P. M., and Rankin, R. E. Effects of incentives upon reactions to peripheral stimuli. *Journal of Experimental Psychology 44,* 400-406, 1952.

Benson, S. K. The effects of monetary rewards on reading level, attitude toward reading and self concept of ability and school achievement. (Doctoral dissertation, University of Southern California, 1970). *Dissertation Abstracts International* 31, 1612-A (1970). (University Microfilms no. 70-19, 108).

Broadhurst, P. L. The interaction of task difficulty and motivation: The Yerkes-Dodson law revived. *Acta Psychologica 16,* 321-338 (1959).

Davis, R. T. and Lovelace, W. E. Variable rewards and peripheral cues in discriminations by irradiated and nonirradiated monkeys. *The Journal of Genetic Psychology 103,* 201-205 (1963).

Deci, E. L. Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology 18,* 105-115 (1971).

Glucksberg, S. The influence of strength of drive on functional fixedness and perceptual recognition. *Journal of Experimental Psychology 63,* 36-41 (1962).

Glucksberg, S. Problem solving: Response competition and the influence of drive. *Psychological Reports 15,* 939-942 (1964).

Haddad, N. F.; McCullers, J. C. and Moran, J. D. Satiation and the detrimental effects of material rewards. *Child Development,* in press.

Johnson, R. and Thompson, C. Incidental and intentional learning under three conditions of motivation. *American Journal of Psychology 75,* 284-288 (1962).

Kruglanski, A. W.; Alon, S., and Lewis T. Retrospective misattribution and task enjoyment. *Journal of Experimental Social Psychology 8,* 493-501 (1972).

Kruglanski, A. W.; Friedman, I. and Zeevi, G. The effects of extrinsic incentive on some qualitative aspects of task performance. *Journal of Personality 39*, 606-617 (1971).

Lepper, M. R. and Greene, D. Turning play into work: Effects of adult serveillance and extrinsic rewards on children's intrinsic. *Journal of Personality and Social Psychology*, in press.

Lepper, M. R.; Green, D. and Nisbett, R. E. Undermining children's intrinsic motivation with extrinsic rewards: A test of the overjustification hypothesis. *Journal of Personality and Social Psychology 28*, 128-137 (1973).

Luchins, A. S. Mechanization in problem solving: The effect of Einstellung. *Psychological Monographs 54*, (6, Whole No. 248) (1942).

McCullers, J. C. and Martin, J.A.G. A reexamination of the role of incentive in children's discrimination learning. *Child Development 42*, 827-837 (1971).

McGraw, K. O. and McCullers, J. C. The distracting effect of material reward: An alternative explanation for the superior performance of reward groups in probability learning. *Journal of Experimental Child Psychology 18*, 149-158 (1974).

McNamara, H. J. and Fisch, R. I. Effect of high and low motivation on two aspects of attention. *Perceptual and Motor Skills 19*, 571-578 (1964).

Miller, L. B. and Estes, B. W. Monetary reward and motivation in discrimination learning. *Journal of Experimental Psychology 61*, 501-504 (1961).

Schere, R. A. Differential reinforcement with exceptional children. (Doctoral dissertation, New York University, 1969). *Dissertation Abstracts International 31*, 1088-A (1970). (University Microfilms no. 70-15, 981).

Spence, J. T. The distracting effect of material reinforcers in the discrimination learning of lower- and middle-class children. *Child Development 41*, 103-111 (1970).

Spence, J. T. and Dunton, M. C. The influence of verbal and nonverbal reinforcement combinations in the discrimination learning of middle- and lower-class preschool children. *Child Development 38*, 1177-1186 (1967).

Spence, J. T. and Segner, L. L. Verbal versus nonverbal reinforcement combinations in the discrimination learning of middle- and lower-class children. *Child Development 38,* 29-28 (1967).

Terrell, G., Jr; Durkin, K. and Wiesley, M. Social class and the nature of the incentives in discrimination learning. *Journal of Abnormal and Social Psychology 59,* 270-272 (1959).

Viesti, C. R., Jr. Effect of monetary rewards on an insight learning task. *Psychonomic Science 23,* 181-182 (1971).

Weick, K. E. Reduction in cognitive dissonance through task enhancement and effort expenditure. *Journal of Abnormal and Social Psychology 68,* 533-539 (1964).

Weinstein, L. Effects of an increment in monetary incentive magnitude on instrumental responding and repeated increases in reward magnitude in humans. *Psychonomic Science 25,* 235-236 (1971a).

Weinstein, L. Effects of a reduction in reward magnitude on active avoidance behavior in humans. *Psychonomic Science 25,* 205-206 (1971b).

Weinstein, L. Negative contrast with humans as a function of emotionality. *The Journal of Psychology 80,* 161-165 (1972).

Yerkes, R. M. and Dodson, J. D. The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative and Neurological Psychology 18,* 459-482 (1908).

FOOTNOTES

1. A water-jar problem is best defined by an example: A mother sends her son to the well to get three quarts of water. She gives him a five-quart can and an eight-quart can. How can the boy get exactly three quarts of water using only these containers and not guessing at the amount?

2. Evidence from Luchins (1942) bolsters this contention. By putting a set breaker in the middle of a series of dual solution problems, Luchins sought to "disrupt the tendency to repeat blindly the B-A-2C method...." In one study from the 1942 monograph where Luchins' data are comparable to data collected here, 84 per cent of the college-age students continued to use the set solution despite their experience with the set breaker.

APPENDIX A

LITERATURE REVIEW

The literature review done in conjunction with the dissertation research was extensive but will not be reproduced in full here because it is to be published in David Greene and Mark R. Lepper (Eds.), *The Hidden Costs of Reward*, Hillsdale, N. J.: Lawrence Erlbaum Associates, book in preparation. This will be a more readily available source than the present and therefore only a synopsis of the review paper appears below.

Section A of the paper reviewed studies which had shown reward to have a detrimental effect on performance and noted that the evidence for a detrimental effect came from studies employing a wide variety of tasks in which a large number of subject and methodological parameters have been varied. In Section B the question of when reward would prove to be detrimental and when facilitating was considered. A comparison of tasks on which there is facilitation with tasks on which there is a detrimental effect seemed to suggest that there are two important dimensions along which a task must be scaled before a prediction for the effect of reward on performance can be made. These were the attractive-aversive and algorithmic-heuristic dimensions. If a task is aversive to the subject then reward can not be expected to disrupt his performance. To the contrary, it will improve it. But it is not sufficient for a task to be attractive to the subject. It must also be designed such that the subject who is superior at heuristic behavior will have a performance advantage. In verbal learning, therefore, the task had to be designed so that subjects could considerably shorten the associative learning process by developing mediating and coding strategies of their own to relate otherwise unrelated items and to differentiate otherwise similar items. In concept formation, the task had to be one employing enough dimensions that breaking down the stimuli into mere concatenations of stimulus values would not be an obvious way to proceed. In the more complex problem solving tasks, there was the similar requirement that direct discovery of the first step for the solution be somehow blocked. In perceptual recognition, the requirement was that subjects neither have nor be given a set for what they are to perceive. In concluding Section B it was noted that incidental learning formed a special case in that the effect of reward on incidental learning was mediated by the offer of reward for intentional learning. Because it may be "incidental learning" which accounts for the superior performance of nonreward subjects on tasks requiring heuristic solutions, it was suggested that determining the reason for reward's effect in the incidental learning paradigm may provide an explanation for reward's detrimental effect elsewhere.

After laying out what appear to be the crucial features for determining when a task will be susceptible to a detrimental effect of reward, it was pointed out that emphasis on these dimensions was not wholly original. Spence (1956) used the position of the correct response in the habit family hierarchy as a determinant of the effect of incentive on performance and that notion covers much of what is meant by making the nature of the solution (algorithmic or heuristic) a determinant of reward's effect. Algorithmic solutions would translate as responses that are readily available at the top of the hierarchy, and heuristic solutions would be those that are initially low in the hierarchy. Spence's prediction, then, that incentive (K) would retard discovery of the correct solution when it was low in the hierarchy but facilitate it when high in the hierarchy is the same as the prediction that heuristic solutions will be slowed but algorithmic solutions facilitated by the offer of reward. If anything important was added to Spence's account it must lie in the inclusion of the aversive-attractive dimension as a second determinant. But again, credit for first focusing on this distinction must lie elsewhere, since Lepper and Greene (1975) used it in explaining when reward would show a detrimental effect on level of motivation. Their distinction was the same only they used the terms "work" and "play" to describe it.

In sum, then, the chief value of the review paper must lie in its integration and reinterpretation of a large number of studies around the central theme of detrimental effects of reward. While provisions for the detrimental effects of reward have been with us for a long time in our theoretical orientations--in addition to Hull-Spence theory there is the Yerkes-Dodson law--detrimental effects appear to have been overlooked in our research emphases. Almost exclusively, research into reward's effects has focused on studies within the aversive-algorithmic task sector.

From the present view this is a mistake since the attractive-heuristic sector is certainly as widely represented by tasks outside the laboratory. An implication which can be drawn from this is that, potentially, reward's detrimental effects are as widespread as its benefits.

APPENDIX B

STATISTICAL TESTS

Experiment 1

One-way ANOVA for Solution Times on Problem Ten

(Reward N=14, Nonreward N=13)

Source	MS	df	F	
Incentive	117,428.57	1	5.58	$p < .05$
Error	21,048.70	25		

Experiment 1

Groups X Trials ANOVA for Solution Times

(Reward N=14, Nonreward N=13)

Source	MS	df	F	p
Between	9925.84	26		
Incentive (A)	57451.00	1	7.16	.01
Error	8024.84	25		
Within	7279.50	243		
Trials (B)	58837.77	9	11.53	.00
A X B	10092.55	9	1.98	.04
Error	5104.65	225		

Experiment 2

Two-way ANOVA for Solution Times on Problem Ten

(Reward N=16, Nonreward N=16)

Source				
Total	67,700.250	31		
Incentive (A)	173,415.188	1	3.4753	.0696
ACT (B)	229,554.625	1	4.6003	.0386
A X B	298,556.000	1	5.9823	.0199
Error	49,899.391	28		

Experiment 2

Repeated Measures ANOVA for Solution Times

(Reward N=16, Nonreward N=16)

Source	MS	df	F	p
Total	16,541.48	319		
Between	25,488.90	31		
Incentive (A)	93,062.06	1	6.3039	.0172
ACT (B)	240,172.31	1	16.2690	.0006
A X B	43,570.81	1	2.9514	.0933
Error	14,762.52	28		
Within	15,578.39	288		
Trials (C)	92,668.44	9	8.5682	.0000
A X C	26,937.28	9	2.4907	.0098
B X C	29,385.84	9	2.7131	.0051
A X B X C	47,519.96	9	4.3938	.0001
Error	10,815.34	252		

APPENDIX C

INSTRUCTIONS TO SUBJECTS

The problems I have for you to work today are called water-jar or water-bucket problems. You may have had problems like these before either as brain twisters or as a type question on an intelligence test. Water-jar problems are best defined by an example, so let me give you one: A mother sends her son to the well to get three quarts of water. She gives him a five-quart and an eight-quart bucket to do that with. How can the boy get exactly three quarts of water using just these containers and not guessing at the amount? The answer is for the boy to fill the larger, eight-quart bucket to the top and then to use the five-quart bucket as a measure to measure out exactly five quarts from the eight-quart bucket. That leaves exactly three quarts in the eight-quart bucket which, of course, is the solution to the problem. Now, in all, you will have ten problems of this general sort to work. Let's begin by looking at some sample problems. (EXPERIMENTER PROCEEDS THROUGH FIVE SAMPLE PROBLEMS GETTING THE SUBJECT TO GIVE HIS ANSWER IN FORMULA FORM AS WILL BE REQUIRED DURING THE TASK.)

Now, during the task I'm going to have you write your answer rather than tell it to me. You should do so on these slips of paper. (BOOKLET OF TEN ANSWER SHEETS, EACH NUMBERED, IS GIVEN TO SUBJECT.) Use this side of each slip for your answer. You may use any of the rest for scratch work. Here's some additional paper should you need it.

When you finish writing your answer and when you are sure it is correct, then tear off the answer sheet and put it in this box along with the problem card. Once you have put the question and your answer in the box, please do not return to either.

Time in this task is not crucial. I will give you as long as you need to find the correct answer to each of the problems. I'm interested only in the solution you find and how you go about finding it. I should stress that each problem does have a solution and if you persist you will find it.

While you are working on the ten problems in this task, I'm going to leave you alone since no one likes to have someone looking over their shoulder while they are trying to solve problems. Consequently I'll go back into the next room. Now from there I can see into here through that one-way mirror, and I'll be checking from time to time to see how you are coming. If during the task you should have any questions just come back and ask me and if, at the end, I fail to notice that you have finished, then just come back and get me, okay? If you have no questions, you may begin whenever you are ready.

REWARD INSTRUCTIONS: Instructions to reward subjects were identical except that just prior to the sentence "If you have no questions, you may begin whenever you are ready", the following instructions were added: "One last thing I should mention is that I've found it helpful if I motivate people a little to get them to work until they find a correct answer on each of the problems. To do that, I'll give you five cents for each correct answer and a fifty-cent (Exp.1)/ one-dollar (Exp. 2) bonus if they are all correct. This means that you will get a dollar/ dollar and a half if all ten problems are correct, 45¢ if nine are correct, and so on. People do much better when rewarded and so, for the time and energy it saves me, it's well worth it. If you have no questions you may begin whenever you are ready. I'll check your answers when you have finished and pay you then."

APPENDIX D

RAW DATA

Experiment 1. Solution Time in Seconds[a] for Subjects in the Nonreward Group. Nonset Solutions Underlined

Subject Category	Problem Number #1	#2	#3	#4	#5	#6	#7	#8	#9	#10
Subjects who broke set only on Problem Ten	114	88	38	40	25	36	26	63	24	<u>215</u>
	100	171	88	132	134	69	83	56	63	<u>216</u>
	80	46	32	39	22	29	40	39	36	<u>138</u>
	106	58	60	80	40	49	69	54	48	<u>137</u>
	65	49	60	71	38	42	34	53	38	<u>63</u>
	84	69	56	69	27	39	37	28	23	<u>38</u>
Subjects who broke set on only one problem prior to Ten and then reverted to B-A-2C.	381	218*	61	48	92	15	45	71	35	<u>27</u>
	59	83	71	108	67	41	67	47	49	<u>92</u>
	49	31	27	34	34	32	20	38	38	<u>15</u>
Subjects who broke set on two or more problems prior to Ten.	63	110	56	68	128	<u>22</u>	<u>16</u>	<u>24</u>	<u>22</u>	<u>17</u>
	99	62	85	82	<u>296</u>	<u>32</u>	<u>20</u>	<u>23</u>	<u>30</u>	<u>19</u>
	199	78	42	100	120	<u>15</u>	126	<u>18</u>	26	<u>40</u>
Subjects who never broke set.	277	113	49	56	47	27	24	16	21	51*
Subjects who never established a set due to errors early in the series	409*	343*	446*	62	62	<u>21</u>	<u>16</u>	<u>09</u>	<u>07</u>	<u>06</u>
	495*	203*	127*	296	150	<u>43</u>	<u>18</u>	330*	140*	<u>21</u>

[a] Errors are starred (*) and nonset solutions are underlined.

Experiment 1. Solution Time in Seconds[a] for Subjects in the Reward Group.

Subject Category	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
	Problem Number									
Subjects who broke set only on Problem Ten	76	62	45	36	48	38*	51*	25*	33	<u>352</u>
	152	71	117	96	123	78	40	34	37	<u>97</u>
	169	95	72	75	68	59	52	78	89	<u>431</u>
	340	50	74	72	48	49	51	68	41	<u>359</u>
	182	418	96	44	38	16	19	25	25	<u>87</u>
	82	76	47	167	42	33	43	44	45	<u>67</u>
	117	74	74	82	304	75	92	96	58	<u>262</u>
Subjects who broke set on only one problem prior to Ten and then reverted to B-A-2C.	207	574*	92	183	80	<u>44</u>	271	91	51	<u>130</u>
	72	91	169*	87	144	<u>66</u>	77	54	88	<u>280</u>
Subjects who broke set on two or more problems prior to Ten.	56	65	39	50	41	30	<u>58</u>	<u>23</u>	<u>75</u>	<u>43</u>
	214*	92	70	77	78	54	<u>27</u>	<u>15</u>	<u>18</u>	<u>09</u>
Subjects who never broke set. (NS=no solution)	162	197	73	64	64	57	61	46	52	NS
	141*	118	38	108	40	40	43	27	64	389*
	77	51	32	48	81*	30	37	56	28	306*
Subjects who never established a set due to errors early in the series.	NS*	126	NS*	342*	250	<u>14</u>	71	214	37	<u>11</u>

[a] Errors are starred (*) and nonset solutions are underlined. NS=no solution, meaning that the subject gave up without finding an answer.

Experiment 2. Solution Time in Seconds[a] and ACT Math Usage Percentage Scores[b] for Subjects in Nonreward Group.

Subject Number	ACT	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
1	.71	87	63	37	42	30	29	30	24	20	116
2	NA	67	132	39	55	41	64	102	58	298	25
3	.69	60	32	30	37	30	25	26	45	34	26
4	.96	99	127	44	164	81	109	104	109	76	19
5	.69	81	107	114	118	81	55	94	63	113	40
6	.30	93	68	62	210	100	36	42	51	34	227
7	NA	33	39	28	69	46	30	39	24	26	13
8	.73	158	40	46	48	52	54	50	45	38	576
9	.89	74	65	87	31	29	28	17	22	20	259
10	.62	400	380	218	352	43	52	30	45	55	342
11	.66	127	58	53	53	122	151	73	68	77	107
12	.48	60	142	67	92	55	75	42	48	47	304
13	NA	174	137	78	128	70	170	89	63	93	88
14	.81	140	76	42	46	25	23	50	29	17	37
15	.62	90	46	52	67	85	81	58	86	62	24
16	.69	78	243	70	106	80	114	77	60	38	89
17	.81	93*	35	76	113	55	56	71	69	47	56
18	.98	34	28	47	79	32	28	27	31	25	36
19	.11	90	454	75	204	137	45	76	60	191	38
20	NA	49	63	55	48	56	23	58	111	41	34

[a] Errors are starred (*)

[b] NA = not available

Experiment 2. Solution Time in Seconds[a] and ACT Math Usage Percentile Scores[b] for Subjects in the Reward Group

Subjects #	ACT	Problem Number #1	#2	#3	#4	#5	#6	#7	#8	#9	#10
1	.40	48	32	163*	48	68	123	258	73	52	1113*(no solution)
2	.99	41*	65	57	111	67	48	94	51	34	113
3	.99	123	45	59	43	46	24	34	42	24	88
4	NA	149	190	67	186	69	60	67	82	68	2400*(no solution)
5	.30	97	162	98	92	187	207	106	89	75	285
6	.11	125	117	87	106	132	97	109	88	45	198
7	.93	79	52	42	51	44	37	63	23	19	54
8	.84	82	43	55	93	46	52	52	84	37	75
9	.61	80	75	56	69	66	87	28	23	20	401*
10	.82	58	96	68	152	51	98	76	72	154	14
11	NA	256	68	81	108	83	77	60	73	30	129
12	.86	110	83	69	124	97	39	43	29	123	370
13	.61	97	99	74	67	169	71	40	50	28	1317
14	.23	176	299	95	59	158	92	59	67	67	986
15	NA	100	130	117	77	107	57	80	164*	75	30
16	NA	60	33	55	53	30	34	34	28	26	62
17	.82	99	84	51	76	66	87	49	36	28	137
18	.17	149	555	126	170	64	208	237	161	116	1680*(no solution)
19	.84	162	242	109	116	120*	146	137	124	74	36
20	.11	120	78	67	191	86	110	219	54	97	48

[a] Errors are starred (*)

[b] NA = not available